



**Elda Gallese**

**Nora Mabel Lac Prugent**

Colaboradora: Lucía Andreozzi

*Instituto de Investigaciones Económicas de la FCEyE de la Universidad Nacional de Rosario*

*Consejo de Investigaciones de la Universidad Nacional de Rosario*

## **PLS PATH MODELING EN LA ALTA COMPLEJIDAD Y ESCASA INFORMACIÓN DEL TERCER MILENIO<sup>1</sup>**

### **1 INTRODUCCIÓN**

El propósito de este artículo es llamar la atención de los estudiantes e investigadores en economía aplicada, sobre las posibilidades de la construcción y evaluación de modelos para tratar con problemas de alta complejidad y escasa información. A la luz de los avances registrados en la implementación de los paquetes de software disponibles y de las múltiples aplicaciones que posibilita -por ejemplo la evaluación de la “satisfacción de los consumidores” europeos (ver Lauro y Vinzi)- parece pertinente hacer una revisión de los mismos. La denominación PLS path model encierra una multiplicidad de modelos. PLS (Partial Least Square) hace referencia al método de estimación, ya que utiliza Mínimos Cuadrados en cada etapa del proceso iterativo (Punto fijo) de la estimación.

El impulso de las transformaciones tecnológicas de los últimos decenios, sin duda incrementó la imperiosa necesidad de invertir, especialmente en Capital Humano (instrucción, formación, capacitación), para activar el desarrollo, asegurar una adecuada formación profesional y crear así nuevos puestos de trabajo para erradicar la pobreza.

En octubre del 2003 el presidente de Argentina presentaba el documento “Objetivos de Desarrollo del Milenio. Argentina La oportunidad para su reencuentro” con estas palabras:

.... Ellos reflejan la factibilidad a partir de los recursos tecnológicos y de conocimiento que la humanidad dispone, de reducir la pobreza extrema y el hambre, así como mejorar la cobertura, calidad y equidad en educación y salud, realizándolo con políticas de desarrollo sostenible y promoviendo valores de equidad y solidaridad: de género, generacionales y territoriales...

Para monitorear si estas metas se están alcanzando es necesario contar con modelos que estimen el Capital Humano en un entorno de alta complejidad y escasa información en cuanto a la distribución de las variables.

El profesor Dagum (1994) explica, el ingreso de las familias en función de la “riqueza neta” (RN) y el “capital humano” (CH) que ellas poseen. En dicho trabajo el profesor Dagum propone que el valor monetario de la *variable latente* capital humano sea estimada por medio del *PLS path modeling*. Es el modelo más simple de los Path ya que cuenta con una sola variable latente y un solo grupo de indicadores, en cierto modo es similar al de *Componentes principales* utilizado por Filmer y Princhett (1998) para estudiar la relación entre la “rique-

---

<sup>1</sup> Trabajo realizado en el marco del Proyecto de Investigación “¿Se están alcanzando los objetivos de desarrollo del milenio?”. Programa 202 de la Secretaría de Ciencia y Tecnología de la Universidad Nacional de Rosario. Dirección: Nora Lac Prugent. [1ECO62]



za” del hogar y la probabilidad de “asistencia a la escuela en los niños” en los estados de la India.

En trabajos posteriores, el profesor Dagum junto con Vittadini y Lovaglio, mejora esta estimación con un modelo, también PLS path model, donde incorpora más variables latentes y más indicadores. En este trabajo usa la denominación de indicadores formativos, porque ellos “forman” o “causan” el constructo multidimensional HC. El Capital Humano es una “consecuencia” de la inversión en educación. Se mide por un conjunto de indicadores formativos. Los indicadores reflectivos, son aquellos que son consecuencia de la cantidad y tipo de inversión en educación.

Interpretando la profunda investigación realizada por Temme, Kreis y Hildebrandt (2006), se puede decir que en los últimos 20 años los modelos de tipo PLS path model no han sido aplicados con frecuencia, aunque sus algoritmos básicos fueron desarrollados en los años setenta y el primer software fue accesible al público en la década del ochenta, por ejemplo el LVPLS (Lohmöller, 1984). El uso restringido del PLS path model en las últimas décadas puede haberse debido en cierto grado a la falta de actualización de los software existentes poco amigables. Es importante destacar que, en la actualidad se puede elegir entre diversos software alternativos, amigables y accesibles, invitando a reflatar y destacar la importancia de *PLS path modeling*. Dichos avances se ponen de manifiesto en la sección dos.

El tercer apartado resume el marco teórico de PLS path modeling de Herman Wold en el contexto de los modelos multiecuacionales. La sección 4 hace referencia a la utilización de dicho modelo en la estimación de “erarquías urbanas”, aplicado a las localidades de la Extensión Metropolitana de Rosario, en el período 1990-1992. (ver Gallese y otros 1992). Finalmente se expondrán algunas reflexiones finales.

## 2 COMPARACIÓN ENTRE LOS DISTINTOS SOFTWARE

*Interpretando y resumiendo la investigación realizada por Temme, Kreis y Hildebrandt (2006),*

El uso restringido del PLS path model en las últimas décadas puede haberse debido, como ya se mencionó, en cierto grado a la falta de actualización de los software poco amigables existentes. En la actualidad esta situación ha cambiado radicalmente, se puede elegir entre diversos software alternativos (*PLS-GUI*, *VisualPLS*, *PLS- Graph*, *SmartPLS*, *SPAD-PLS*) los cuales presentan un adelanto en facilitar la interacción con el usuario entre otras cosas. Dichos programas contienen en mayor o en menor medida implementaciones de los algoritmos desarrollados por Wold (1982) y Lohmoller (1984). El PLS path model evita el problema de la *identificación*. Es importante remarcar que todos estos programas están en constante desarrollo. A continuación se presenta una breve descripción.

*LVPLS* (1984, 1987). El programa LVPLS 1.8, corre bajo DOS, incluye dos módulos para la estimación de Path models. LVPLSC analiza la matriz de covariancia de las variables observadas, y LVPLSX procesa la matriz de datos (row data). Cuando trabaja con la matriz de datos (row data) ofrece blindfolding y jackknifing como métodos de remuestreo. En caso de usar la matriz de covariancia/correlación como input, no.

*PLS-GUI* (2005). Este programa corre en un entorno de Windows. Es amigable para el input conduciendo al usuario paso a paso. La salida es la misma que LVPLS. La versión actual ofrece la opción de bootstrap, no incluida en LVPLS.

*VisualPLS* (2006). Últimamente permite el análisis de la matriz de datos (diferentes formatos de datos) solamente. Corre bajo un entorno de Windows. La especificación del modelo se realiza marcando las variables latentes y asignando los indicadores en una ventana **pop-up**. Además de blindfolding y jackknifing ha sido integrado la opción de bootstrapping. Tiene un



soporte especial para especificar efectos moderados y un factor de segundo orden.

*PLS-Graph* (2003). Permite el análisis de la matriz de datos solamente. Corre bajo un entorno de Windows. En cuanto al momento de especificar el modelo, presenta una interfase gráfica provee algunas herramientas para dibujar el diagrama. Las estimaciones resultantes se presentan en formato ASCII y también como en un gráfico de path model. Los métodos de resampling incluyen blindfolding, jackknifing y bootstrapping.

*SmartPLS* (2005). Este programa es independiente del sistema operativo ya que está basado en Java. Solo trabaja con matriz de datos. Se especifica el modelo para estimar el modelo estructural para las variables latentes asignando los indicadores de las variables latentes. Suministra el output en formato HTML, Excel o LaTeX. Los métodos de remuestreo accesibles son, bootstrapping y blindfolding.

Tiene una rutina para detectar heterogeneidad, la "finite mixture routine" (FIMIX). Dicha opción es de interés si una heterogeneidad no observada en los datos es esperada.

*SPAD-PLS*: Es parte del software para análisis de datos SPAD. Corre bajo Windows. Necesita como input la matriz de datos. Se puede especificar un modelo con un menú o gráficamente en Java applet. Tiene diferentes opciones para manejar datos perdidos y estima la multicolinealidad. Los resultados son registrados como un path diagrama y como texto o archivo de Excel. Son accesibles blindfolding, jackknifing y bootstrapping, incluyendo intervalos de confianza.

## 2.1 Datos Faltantes

En cuanto a tratar con datos faltantes, se han propuesto varias alternativas. LVPLS ofrece un tratamiento específico que combina imputación del valor medio y **deletion** apareada. El tratamiento de datos faltantes es también provisto por VisualPLS y PLS-GUI, así también como en PLS-Graph y SPAD-PLS. En cambio SmartPLS ofrece dos opciones: sustituye los valores faltantes por la media calculada a través de todos los casos disponibles de una variable, o borra los casos con datos faltantes (casewise deletion), este último procedimiento desperdicia una gran cantidad de información y conduce a una menor eficiencia, por ello no es recomendable.

## 2.2 Multicolinealidad

La Multicolinealidad puede ser un problema tanto en la estimación de los pesos de los indicadores como en la estimación en la relación entre las variables latentes. SPAD-PLS es en la actualidad el único programa que está teniendo en cuenta este problema y para ello, provee una rutina de regresión PLS para abordar este tema.

## 2.3 Métodos de remuestreo

PLS-Graph y SmartPLS permiten al usuario elegir entre dos procedimientos de corrección: en la primera (cambio de signos individual) el signo de cada peso individual **outer** se iguala al correspondiente signo en la muestra original. Esta opción debe ser usada con cuidado, ya que no apela a la coherencia. La segunda opción (**construct levels changes**) compara las cargas para cada variable latente con las originales y revierte el signo de los pesos si el valor absoluto de la diferencia sumada entre a carga original y la del bootstrap es mayor que el valor absoluto de la suma de las cargas originales y las bootstrap. Igualmente ambos procedimientos no aseguran que el cambio de signos sea correctamente manejado. PLS-GUI y VisualPLS solo presentan la segunda opción. Desde que SPAD-PLS utiliza los elementos del primer autovector de un análisis de componentes principales, con predominio de signos positivos, el control de signos alinea los signos de las muestras del bootstrap, a aquellos en



la muestra original.

Otra manera de analizar la significación de las estimaciones PLS son los intervalos de confianza a partir de las muestras bootstrap, esta opción, utilizando el método de percentiles aparece únicamente en SPAD-PLS.

## 2.4 Otras Particularidades

Con respecto a las cargas internas todos los programas ofrecen esquemas para estimar el **inner model** (centroide, factor path weighting) ya disponible en LVPLS. Un tema de especial interés es el uso de diferentes conjuntos de valores iniciales para determinar los outer weight. Los valores iniciales pueden tener un impacto en el signo de los pesos estimados o las cargas de los factores, y más aun en los coeficientes del path model. Aunque esta o es una cuestión estadística, es importante a la hora de interpretar los resultados. Ninguno de los programas permite al usuario elegir y especificar los valores iniciales.

## 2.5 Facilidad de Uso

Comparado a LVPLS, todos los software recientes son considerablemente más amigables, mas aun cuando el programa permite al usuario especificar el modelo gráficamente y además un diagrama del modelo en la salida. En PLS-Graph, SPAD-PLS y SmartPLS es fácil cambiar el conjunto de datos sin tener que especificar nuevamente el modelo. Adicionalmente puede guardarse el análisis completo, data set, modelo y resultados en un único archivo. La mayoría de los programas proveen variadas cajas de herramientas que ayudan a mejorar el layout de los diagramas, (color, tamaño, texto, etc.), especialmente en el caso de SPAD-PLS y SmartPLS. Para todos los programas los manuales del el usuario documentan aplicaciones con datos ejemplo.

## 2.6 Características generales

Comparando el software en cuanto a sus requerimientos del sistema, es claro que usuarios de UNIX/LINUX o Mac deben usar SmartPLS ya que es único que posee una plataforma independiente. Todos los programas en la actualidad están diseñados para variables latentes continuas u categóricas ordinales con cinco o más categorías de respuesta. Además variables binarias exógenas pueden incluirse en el análisis. Si solo se cuenta con matrices de covariancia como datos para ingresar, la elección esta únicamente restringida a LVPLS o PLS-GUI. Excepto por LVPLS todos los programas requieren una definición común de los valores faltantes para todas las variables. En general, todos los programas son capaces de procesar datos en formato ASCII, y algunos requieren una conversión a un formato específico de datos. SPAD-PLS también soporta formatos de datos de paquetes como SPSS y SAS.

LVPLS y los software más recientes han sido comparados y caracterizados. Se ha puesto énfasis en la facilidad de uso y las opciones metodológicas. Los nuevos programas han dado un gran paso en cuanto a la facilidad de uso. Puntos especiales se detectan en SmartPLS y VisualPLS que permiten la estimación de efectos de interacción y útiles opciones para exportar en el caso de SmartPLS y SPAD-PLS.

Un avance metodológico es el procedimiento bootstrap, para testar la significación de los parámetros estimados, que esta presente en todos los paquetes y reemplaza las rutinas blindfolding y jackknifing de remuestreo presentes en LVPLS, mas aún SPAD-PLS presenta estimación bootstrap de intervalos de confianza para los parámetros. Otro aspecto importante: la validación del modelo, como algunas medidas como el índice de bondad de ajuste, en este caso solamente se presentan los índices de validación cruzada del blindfolding.

## 3 PLS PATH MODEL

Al enfrentarse con un problema complejo el primer paso, fundamental, es la formulación del

modelo. Ello se efectúa teniendo en cuenta ciertos supuestos “a priori”, que parten de la observación de la realidad, de la experiencia anterior, de los elementos que proporciona la teoría del tema en cuestión, etc. Todo depende en última instancia de la sagacidad y buen sentido de aquel que está realizando la investigación y se ocupará de la formulación de los modelos tratando de interpretar el fenómeno de la forma más fidedigna posible, con la mayor aproximación, tratando de minimizar los errores y aumentar la precisión de su modelo.

El PLS path model fue diseñado por H. Wold (1980) principalmente para el análisis causal-predictivo, de problemas con alta complejidad y poca información.

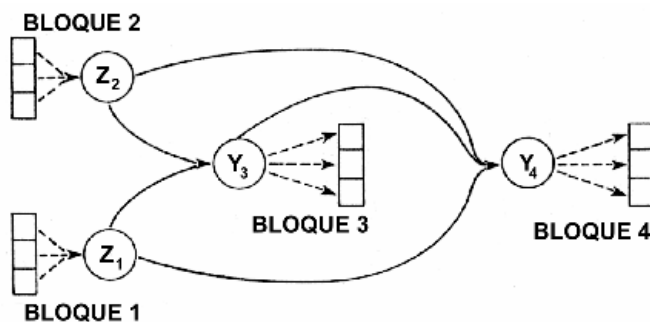
La *alta complejidad* está dada por la intervención de un gran número de indicadores (variables observables), relacionados a variables no observables o variables latentes.

La falta de relevamiento de datos a través del tiempo, y por consiguiente el desconocimiento de las propiedades de las distribuciones estadísticas de las variables relevantes, reflejan el hecho de *poca información y escaso conocimiento*.

### 3.1 Definición teórica del modelo

Hay dos tipos de indicadores o variables manifiestas (MV): indicadores formativos e indicadores reflectivos. Los indicadores reflectivos son variables observables, las cuales, se supone, están afectadas por la misma variable latente; los indicadores formativos son variables observables, las cuales, se supone, forman o causan cambios en la variable latente. La especificación de formativo o reflectivo depende de la prioridad causal entre el indicador y la variable latente. En el siguiente esquema, los indicadores vinculados con las variables exógenas (Z-LV) son formativos, aquellos vinculados con las variables endógenas (Y-LV) son reflectivos.

El esquema de flechas ilustra el modelo teórico conceptual. En él se observan:



- Cuáles son las variables latentes, **LV** que entran en el modelo.

Las variables (LV) exógenas:  $Z_1$  y  $Z_2$

Las variables (LV) endógenas  $Y_3$  e  $Y_4$

- Cuáles son los indicadores (variables observables o manifiestas MV), que intervienen en cada bloque para medir indirectamente a las variables latentes.

- Qué relaciones internas se supone que existen entre las variables latentes.

Una variable latente LV, combina los ítems de un bloque (indicadores, MV), para ser usado luego en el modelo causal – predictivo.

Notación:

$J = 4$  : número de bloques o variables latentes



$K_j$  = número de indicadores en el bloque j-ésimo  
 $Z_1$  y  $Z_2$  = variables latentes predeterminadas (o exógenas)  
 $Y_3$  e  $Y_4$  = variable latente endógena  
 $x_{jk}$  = k-ésima variable observable (MV) del j-ésimo bloque

- La relación de cada indicador con su bloque (LV) configura la ESTRUCTURA DE BLOQUE, ilustrada en el diagrama por las líneas de puntos.

Formalmente:

$$\begin{aligned} x_{jk} &= f(Z_j) + v_{jk} & j &= 1, 2 \\ x_{sk} &= f(y_s) + v_{sk} & s &= 3, 4 \end{aligned} \quad (1)$$

con

$$\text{var}(Z_j) = \text{var}(y_s) = 1$$

$$r(Z_j Z_i) = \rho_{ij} \neq 0$$

$$r(y_s y_r) = \rho_{sr} \neq 0$$

$$r(v_{ih}, v_{ik}) = r(v_{ih} Z_i) = r(v_{ih} y_j) = 0$$

$$i \neq j \quad i, j = 1, \dots, j \quad h = 1, \dots, K_i \quad k = 1, \dots, K_j$$

- Las relaciones de las variables latentes entre sí, son las RELACIONES INTERNAS (sistema de ecuaciones interdependientes)<sup>2</sup>, ilustradas en el diagrama por flechas de líneas llenas.

- Las flechas, indican los canales de información del modelo, mostrando cuáles variables son exógenas ( $Z_1$  y  $Z_2$ ), y cuáles son endógenas ( $y_3$  e  $y_4$ ).

Formalmente:

$$y_3 = \alpha_3 + \gamma_{31} Z_1 + \gamma_{32} Z_2 + \delta_3 \quad (2)$$

$$y_4 = \alpha_4 + \beta_{43} y_3 + \gamma_{41} Z_1 + \gamma_{42} Z_2 + \delta_4$$

- La especificación de las RELACIONES INTERNAS y la ESTRUCTURA DE BLOQUE, constituye la "DEFINICIÓN TEÓRICA DEL MODELO".

- Una característica del modelo que no está ilustrada explícitamente por el esquema de flechas, es la especificación de las RELACIONES CAUSAL-PREDICTIVA.

- La ESTRUCTURA DE BLOQUE, las RELACIONES INTERNAS y las RELACIONES CAUSAL – PREDICTIVA, se denominan "RELACIONES ESTRUCTURALES" del modelo.

*Es pertinente recordar, que las relaciones causales son siempre predictivas, pero las relaciones predictivas no son necesariamente causales. La decisión respecto de la causalidad de una variable exógena es del especialista del tema en estudio. No es responsabilidad de la estadística.*

### 3.2 Consideraciones generales.

En las relaciones entre las variables latentes y sus indicadores, el modelo da dos medidas de la relevancia de un indicador, su ponderación y su carga.

- La carga de un indicador, mide la relación individual entre el indicador y la variable latente de su propio bloque.

- La ponderación de un indicador, por su ubicación en la construcción de la variable latente, mide la relación entre el indicador y su propio bloque, removiendo la influencia de

<sup>2</sup> Para una introducción a los modelos multiecuacionales (sistema de ecuaciones interdependientes) consultar Gallese (1998)

los otros indicadores del mismo bloque.

Por el modo de obtención, la ponderación de un indicador también mide el grado de relación de éste con las variables latentes de los bloques que están vinculados con él por medio del diagrama de flechas.

Las dos medidas son de importancia si se las interpreta correctamente, por consiguiente en los resultados, se presentan ambas.

Para el ajuste del modelo se recomienda ver Wold (1980), Bookstein (1980) y Lohmöller (1984, 1987). Los algoritmos básicos utilizados por los paquetes modernos de procesamiento de *PLS path modeling* se encuentran en estas publicaciones.

#### 4 MODELO DE JERARQUIAS URBANAS

Se ilustra el Path model que en el año 1990, en la Subsecretaría de Planeamiento de la Provincia de Santa Fe, se construyó para:

-Ordenar jerárquicamente las quince localidades que conforman el “área de estudio”<sup>3</sup> teniendo en cuenta su “nivel urbano”<sup>4</sup>

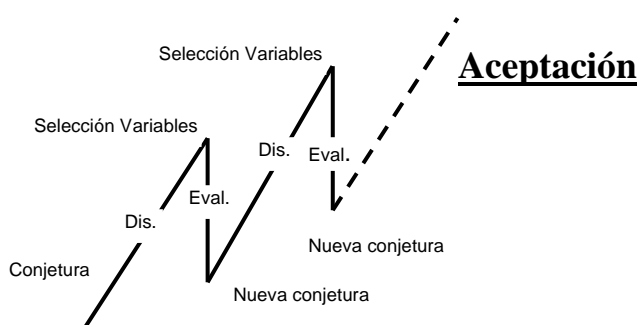
-Diseñar un modelo que explique el “nivel urbano” en términos de condiciones “sociales, económicas y urbanísticas”.

Para su desarrollo, debido a la complejidad del “fenómeno” abordado y a los objetivos de construir “proyecciones” sobre el área en múltiples aspectos (políticos, socioeconómicos, urbanísticos, ambientales, etc.), se conforma un equipo de carácter pluridisciplinario: cinco arquitectos urbanistas, una estadística, una licenciada en Economía, una antropóloga social, y una licenciada en Ciencia Política.

En este marco, la aplicación del modelo tentativo de jerarquías urbanas (que a continuación se explicita) tiende en particular hacia la definición de los roles que cada centro poblacional cumple en el juego de interdependencias del área de estudio.

##### 4.1 Construcción del modelo

La construcción del modelo es un proceso iterativo y de evolución. A partir de los objetivos, se intenta un primer diseño del modelo que conduce a la selección de variables tanto observables (Variables Manifiestas) como no observables (Variables Latentes LV).



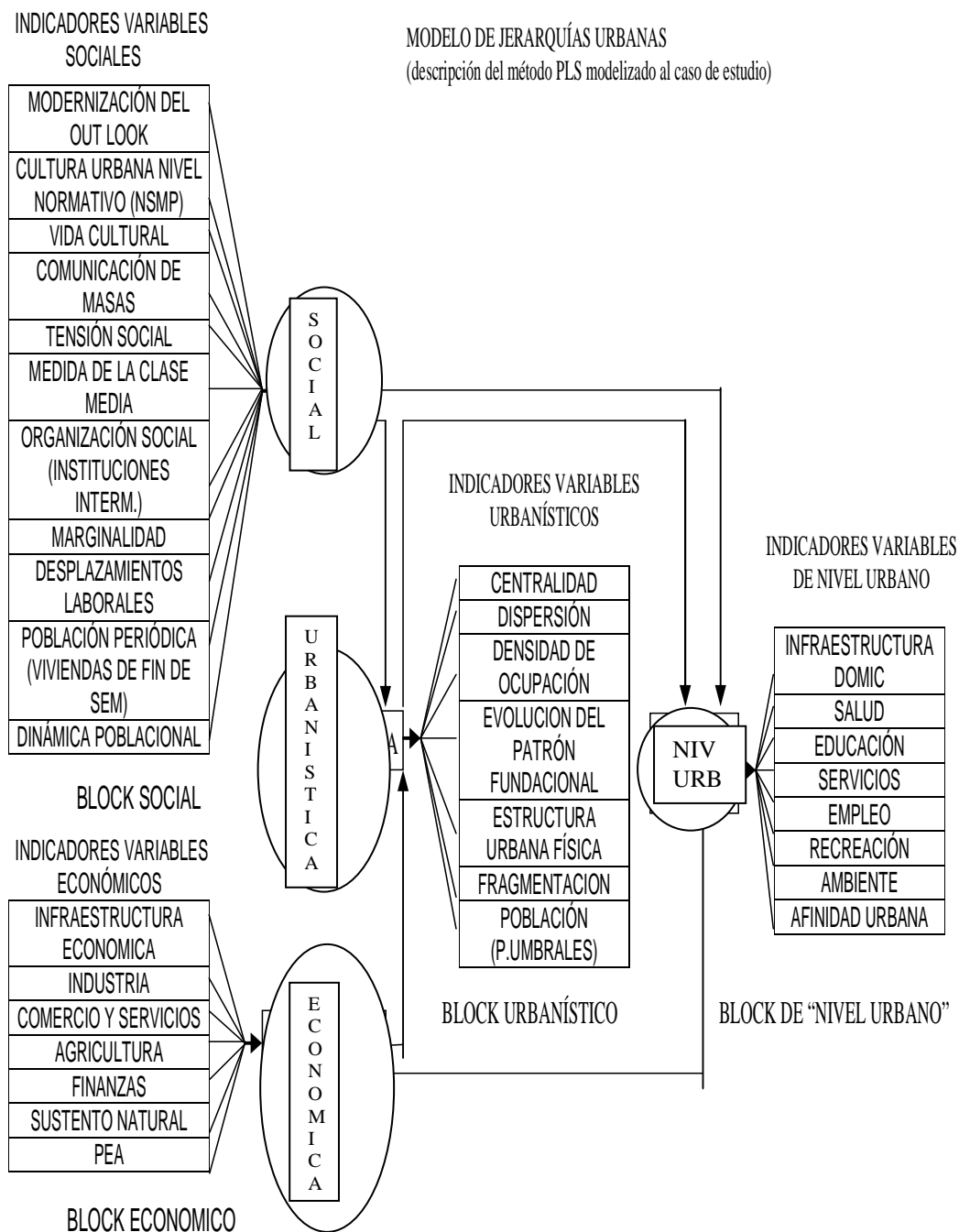
El análisis y la discusión de este diseño conducen a una nueva conjetura y así, sucesiva-

<sup>3</sup> A los efectos de dimensionar la aplicación del “modelo”, dicha área de estudio se correspondería con lo que la investigación ha definido como “extensión metropolitana”.

<sup>4</sup> La noción de “nivel urbano” que constituye la variable decisiva que cierra el “modelo” se corresponde, según lo acordado en la discusión interdisciplinaria, con la cantidad, calidad y diversidad de equipamientos y servicios que brinda a su población cada centro urbano considerado, “aunque dicha oferta se ha ampliado a algunos indicadores no convencionales que se remiten a condiciones vivenciales y ambientales que nos aproximaría a lo que la interpretación sociológica define como “calidad de vida”, o la ecológica como “calidad ambiental” o “ambiente urbano”.

mente, ilustrado en el esquema anterior.  
Esta etapa condujo al siguiente modelo:

### Esquema del “modelo de jerarquías urbanas”.



El conocimiento teórico-conceptual contenido en el modelo está ilustrado en el “esquema de flecha”.

Las nociones elementales son las siguientes: las condiciones sociales y económicas afectan





tanto a la condición urbanística como al nivel urbano. A su vez la condición urbanística afecta claramente al nivel urbano.

Las variables observables que son usadas como indicadores de los cuatro bloques, se muestran en su carácter de formativas (Bloques Social y Económico) y reflectivas (Bloques Urbanístico y Nivel urbano).

El propósito predictivo del modelo respondería “tentativa e inicialmente” a medir el efecto “relativo” de cada una de las condiciones (social, económica y urbanística) sobre el nivel urbano.

El esquema de Flechas es el plano teórico-conceptual del modelo. En él se indica cuáles son las LV que entran en el modelo, que indicadores observables serán usados y su agrupamiento dentro del block, y qué “relaciones internas” se supone que existen entre las LV. El investigador, junto con los especialistas en un trabajo interdisciplinario, diseña el esquema de flechas sobre la base conjunta de su conocimiento, su experiencia y su intuición acerca del modelo explorado y los datos que están a su disposición.

Las LV son la unidad estructural del modelo y están relacionadas por “relaciones internas”. La relación formal entre las LV y sus indicadores constituye la “estructura de bloque”. Las flechas indican canales de información en el modelo.

#### Definición de variables

Cada variable seleccionada en el diseño del modelo debe ser definida con precisión y objetividad.

### 4.2 Ajuste del modelo y estimación de los parámetros con PLS Path model

Se utilizó el método de los mínimos cuadrados parciales (Partial Least Squares PLS). Este método fue diseñado por Herman Wold para analizar problemas de alta complejidad y poca información.

Siendo a distribución libre, la estimación PLS no impone ninguna restricción sobre el formato de los datos. Los datos disponibles pueden ser series temporales, o de corte transversal. Las observaciones de los indicadores pueden ser mediciones cuantitativas, ranqueo ordinal, o registros de ocurrencias, no-ocurrencias, etc.

El contenido empírico se extrae de los datos, y el modelo es mejorado a través del procedimiento de estimación, por la interacción entre el modelo y los datos y la reacción del investigador. El uso de muchas variables observables enriquece el contenido empírico del modelo y aumenta la precisión de la estimación PLS. En la interacción entre los datos y el modelo original aparecerá cuáles indicadores son relevantes y cuáles pueden ser omitidos.

#### Procesamiento de los datos

Ningún paquete disponible resultó útil y el procesamiento lo hizo el Licenciado en Estadística Daniel Wojdyla programando paso a paso con el software SAS.

#### Interpretación de los resultados

Si el análisis confirmatorio es satisfactorio se podrá dar respuesta a varias cuestiones: se tendrán indicadores del nivel urbano de cada localidad, lo que permitirá ordenar jerárquicamente las quince localidades;

- se podrá evaluar la contribución “individual” de un indicador a la relevancia de la variable latente correspondiente a su bloque;
- se podrá también medir la contribución “relativa” de un indicador a la relevancia de la variable latente correspondiente a su bloque;
- se podrá explicar el comportamiento del nivel urbano en términos de las condiciones sociales, económicas y urbanísticas de cada localidad;



- se podrán realizar predicciones a través de las relaciones causal-predictivas. Recordemos que las relaciones causales son siempre predictivas, pero las relaciones predictivas no son necesariamente causales. La causalidad debe sustentarse desde otro lugar; la causalidad no es asunto de la estadística.

Si el análisis confirmatorio no es satisfactorio, habrá que indagar las fuentes del problema y subsanarlo.

Este modelo, por su flexibilidad se ha aplicado al análisis de problemas pertenecientes a distintas disciplinas. Por ejemplo:

En el área de Educación y Sociología se usa para intentar explicar el “nivel alcanzado por los alumnos en términos de condiciones en el hogar y condiciones en la escuela”.

En el área de Economía y Sociología se usa para intentar y explicar la “cantidad y calidad de las oportunidades laborales en términos de productividad, demanda de trabajo y oferta de trabajo”.

En el área de Economía, Sociología y Ciencias Políticas se usa para intentar explicar “el crecimiento económico en términos de condiciones sociales, económicas y políticas”.

Este tipo de modelo “soft” se usa y perfecciona en instituciones como el “Centro para el crecimiento y desarrollo humano” y el “Instituto de Investigaciones Sociales” de la Universidad de Michigan, y en el “Centro de Investigaciones de Economía Matemática para la Planificación de París”, entre otras.

La implementación y ajuste del modelo demandó un esfuerzo interdisciplinario de casi tres años y los resultados fueron satisfactorios.

## 5 REFLEXIONES FINALES

La adopción del PLS Path model se justifica a nivel teórico por su capacidad predictiva y por una modelización, se puede decir, más blanda con respecto a las técnicas basadas en el criterio de máxima verosimilitud, y en el aspecto operativo por una mayor flexibilidad en la especificación del modelo.

Es de esperar que el recorrido por el marco teórico del PLS Path model, la aplicación realizada en el problema de Jerarquías Urbanas y sobretodo, la información brindada sobre los nuevos paquetes de procesamiento de datos, accesibles y amigables, sirvan de estímulo a los jóvenes investigadores para modelar y evaluar los problemas de alta complejidad cuando se cuenta con poca información sobre la distribución de las variables.

## REFERENCIAS BIBLIOGRÁFICAS

- Bookstein, F. L. (1980). Data Analysis by Partial least Squares. *Evaluation of Econometrics Models*. Edit by KMENTA, J. And RAMSEY, J. B Academic Press. Page. 75-90.
- Dagum, C. (1994). Human Capital, Income and Wealth Distribution Model With Applications. S.A.E. Cuaderno N° 11.
- Filmer, D. y Pritchett, L. (1998). Estimating Wealth effect Without Expenditure Data. With an Application to Educational Enrollments in States of India. *Policy Research Working Paper*. The World. Bank.
- Gallese, E., Caballero, A. y otros (1992). Modelo De Jerarquías Urbanas Aplicado A Las Localidades De La Extensión Metropolitana De Rosario. *Actas del Noveno Congreso*



*Nacional de profesionales en Ciencias Económicas.*

- Gallese, E.(1994).Importance of latent path modeling whit partial least squares. A new programme of Econometrics for future Statisticians at Rosario's National University. *Proceeding of Fourth International Conference in Teaching Statistics*. Marruecos,(Volumen 2.), 444.
- Gallese, E. (1998). Capital Humano y PLS Path Model. *Actas de la Tercera Jornadas Investigaciones de la Facultad de Ciencias Económicas y Estadística*. Octubre 1998.
- Lauro, C. y Vinzi, V. E.. *Some contributions to PLS Path Modeling and a system for the European Customer Satisfaction*.  
<http://www.sis-statistica.it/files/pdf/atti/RSMi0602p201-210.pdf> (visitado 10/09/2007).
- Temme, D.; Kreis, H. and Hildebrandt, L. (2006). PLS Path Modeling-A Software Review. *Economic Risk*. Berlin. SFB 649.
- Vittadini, G., Dagum, C. y Lovaglio, P.G.. *A Method For The Estimation Of The Distribution Of Human Capital From Sample Survey Of Income And Wealth*.  
[http://www.csmb.unimo.it/epub/soattiva/doc/costa\\_dagum\\_rel.pdf](http://www.csmb.unimo.it/epub/soattiva/doc/costa_dagum_rel.pdf) (visitado 11/09/2007).
- Lohmöller, J. (1984). *LVPLS Latent Variables Path Analysis With Partial Least Squares Estimation*. Program Manual. Universitatzu Koln.
- Wold, H. (1980). *Model Construction and Evaluation When Theoretical Knowledge is Scarce. Evaluation of Econometrics Models*. Edit by KMENTA, J. And RAMSEY, J. B. Academic Press. Pág. 47-74.
- Presidencia de la Nación (2003). *Objetivos de Desarrollo del Milenio. Argentina. La oportunidad para su reencuentro*.